MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2017/2018

DPS5018 –INTRODUCTION TO PROBABILITY AND STATISTICS

(For Diploma students only)

2 MARCH 2018
9.00 a.m - 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **8** pages (**4** pages with **5** questions and **4** pages for appendix)
2. Answer **ALL** questions. All necessary working steps must be shown.
3. Write all your answers in the answer booklet provided.

QUESTION 1 [15 Marks]

(a) A company claims that they will sent the parcels by Post Express, Skynet and Nationwide Express at 30%, 33% and 37% respectively and the probability that parcels are guaranteed to reach its destination on the following day are 0.98, 0.81 and 0.99 respectively.

i. Construct and label the tree diagram to represent the information above. (4 marks)

ii. Determine the probability that the parcel will reach its destination on the following day. (3 marks)

iii. Given that the parcels will NOT reach the destination on the following day, calculate the probability of the parcels are sent by Skynet. (4 marks)

(b) The car trader has recorded their sale from July to September 2017 as follows.

Sales	July	August	September
National	20	7	28
Foreign	12	6	11

i. Calculate the percentage of national car which sold in September 2017. (2 marks)

ii. Determine the probability of foreign cars being sold or the cars sold in August 2017. (2 marks)

QUESTION 2 [25 Marks]

(a) The table below recorded the ages of a group of 40 women when they get married.

Ages (years old), x	27	29	30	31	32	35
$P(X = x)$	0.125	0.375	0.325	0.0500	0.100	0.025

Calculate the mean and standard deviation of the age. (5 marks)

(b) The management of XeDer Condominium claims that 70% of the cars park in inner parking lot are local cars. If a sample of 10 cars is chosen at random, calculate the probability that

i. all cars are the local cars. (2 marks)

ii. at least 7 cars are local cars. (2 marks)

Continued...

(c) In a sample of 8 working adults is chosen at random, it is found that 40% of working adults do not like their current job.

i. How many of them do not like their current job? (1 mark)

ii. Calculate the probability that less than 6 working adults do not like their current jobs. (2 marks)

(d) A survey had been done on ability of tongue rolling among the students in their classes. They found that on average, 4 students are capable of tongue rolling in each class. Calculate the probability that

i. at most 2 students are capable of tongue rolling in each class. (2 marks)

ii. more than 9 students are capable of tongue rolling in 2 classes. (4 marks)

(e) A model company is doing a survey on the masses of a group of fashion show female models. The masses of the models have a normal distribution with a mean of 56kg and a standard deviation of 15kg. If the body mass is more than 70kg, it is classified as overweight, whereas if body mass is less than 45kg, it is classified as underweight. Hence, the ideal mass for a female model is between 45kg and 70kg. If a model is chosen at random from the group, find the probability of that the model

i. has an ideal mass. (5 marks)

ii. has mass less than 62kg. (2 marks)

QUESTION 3 [10 Marks]

(a) In a sample of 23 cylinders which is being selected randomly, the mean and standard deviation of thickness of cylinders produced by a machine are 9.2cm and 4.0cm respectively. Construct a 95% confidence interval for the population mean. (4 marks)

(b) In a lucky draw organized by AEON Ayer Keroh, 28 out of 500 of customers won RM50 voucher on that day.

i. Find the point estimate for proportion of winning the voucher. (1 mark)

ii. Construct a 90% confidence interval for the population proportion of those who won the RM50 voucher. (5 marks)

Continued...

QUESTION 4 [30 Marks]

(a) Nowadays, majority of Malaysians are facing stress everyday due to the demands of work, family and others. Based on a sample of 180 adults which is being selected randomly, 57% of them are facing this problem.

i. State the null and alternative hypothesis. (2 marks)

ii. By using the hypothesis testing of one population proportion, calculate the test statistic. (2 marks)

iii. Is there any evidence to reject the claim? Test at 5% level of significance. (4 marks)

(b) According to Malaysia Qualification Accreditation (MQA), students need to study about 120 hours for DPS5018 Introduction to Probability and Statistics subject in one semester. To verify whether students have achieved this requirement, the subject lecturer randomly selected a sample of 15 students and found that the sample mean and standard deviation are 110 hours and 12.4 hours respectively. Test at 10% level of significance. (9 marks)

(c) Teacher Harry claimed that the failure rate of Introduction to Probability and Statistic's subject has been reduced in Semester 2 when compared to semester 1 due to quality of students and teaching methodology. To verify this claim, he randomly selected a sample of students from each group and recorded the number of students who failed this subject as follows.

	Semester 1	Semester 2
Number of students who fail, x	35	27
Number of students being selected, n	70	90

i. Determine the sample proportion who fails the subject in these two semester. (2 marks)

ii. State the null and alternative hypotheses. (2 marks)

iii. By using the hypothesis testing of two population proportion, compute the test statistic, Z. (5 marks)

iv. Test the claim at 1% level of significance. (4 marks)

Continued...

QUESTION 5 [20 Marks]

(a) A sale person recorded the mileage (km) and petrol expenditure (RM) for each trip in her notebook as follows.

Mileage (km), x	**Petrol Cost (RM), y**
250	50
281	49
179	39
156	35
120	32
167	34
136	30
234	42
m	45
198	37

i. Given that the total mileage (km) travelled in 10 trips is 1931km, find the value of m. (2 marks)

ii. Refer to table above, calculate the total cost for petrol which spent by the sale person. (1 mark)

iii. Calculate the $\sum y$ and SS_{yy}. (3 marks)

iv. Given that the $SS_{XX} = 23806.9$ and $SS_{XY} = 3023.7$, compute the coefficient of correlation, r. Then, comment your result. (3 marks)

v. Compute b_1 and b_0 to construct the estimated least square regression equation. (5 marks)

vi. Based on this record, if the sale person travels further with the distance of 300km, how much the cost of petrol? (2 marks)

(b) Refer to part (a), the sale person would like to test whether the β_1 is positive at 5% level of significance.

i. State the null and alternative hypothesis. (2 marks)

ii. Given the $s_b = 0.01716$, compute the test statistic. (2 marks)

End of Page.

APPENDIX – KEY FORMULA

- **Mean for ungrouped data**

$$\bar{x} = \frac{\sum x_i}{n} \qquad \text{or} \qquad \mu = \frac{\sum x_i}{N}$$

- **Mean for grouped data**

$$\bar{x} = \frac{\sum x_i f_i}{n} \qquad \text{or} \qquad \mu = \frac{\sum x_i f_i}{N}$$

Where x is the midpoint and f is the frequency of the class

- **Variance for ungrouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \qquad \text{or} \qquad s^2 = \frac{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}}{n-1}$$

- **Variance for grouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{n-1} \qquad \text{or} \qquad s^2 = \frac{\sum x_i^2 f_i - \frac{\left(\sum x_i f_i\right)^2}{n}}{n-1}$$

Where x is the midpoint and f is the frequency of the class

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

- Mean of a discrete random variable x : $\mu = \sum xP(X = x)$

- Standard deviation of a discrete random variable x : $\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$

- Binomial probability formula : $P(X = x) = \binom{n}{x} p^x q^{n-x}$

- Poisson probability formula : $P(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$

- Key formula for Binomial Probability Distribution:
  - If $p \leq 0.5$
    - i. $P(X = r) = B(r) - B(r-1)$
    - ii. $P(X \geq r) = 1 - B(r-1)$
    - iii. $P(X \leq r) = B(r)$
    - iv. $P(X > r) = 1 - B(r)$
    - v. $P(X < r) = B(r-1)$
    - vi. $P(a \leq X \leq b) = B(b) - B(a-1)$
  - If $p > 0.5$
    - i. $P(X = r) = B(n-r) - B(n-r-1)$
    - ii. $P(X \geq r) = B(n-r)$
    - iii. $P(X \leq r) = 1 - B(n-r-1)$
    - iv. $P(a \leq X \leq b) = B(n-a) - B(n-b-1)$

- Key formula for Poisson Probability Distribution:
  - If $\mu = \lambda$
    - i. $P(X = r) = Poi(r) - Poi(r-1)$
    - ii. $P(X \geq r) = 1 - Poi(r-1)$
    - iii. $P(X \leq r) = Poi(r)$
    - iv. $P(X > r) = 1 - Poi(r)$
    - v. $P(X < r) = Poi(r-1)$
    - vi. $P(a \leq X \leq b) = Poi(b) - Poi(a-1)$

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \dfrac{x - \mu}{\sigma}$

ESTIMATION

- The $(1-\alpha)100\%$ confidence interval for μ is

$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

$\bar{x} \pm t_{\alpha/2,n-1}s_{\bar{x}}$ where $s_{\bar{x}} = \dfrac{s}{\sqrt{n}}$

- The $(1-\alpha)100\%$ confidence interval for p is

$\hat{p} \pm z_{\alpha/2}s_{\hat{p}}$ where $s_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

- **Sample size**

$$n = \left[\frac{\left(Z_{\alpha/2}\right)\sigma}{E}\right]^2 \qquad n = \frac{\left(Z_{\alpha/2}\right)^2 pq}{E^2}$$

HYPOTHESIS TESTING

- **ONE population**

Mean:

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \textbf{or} \quad t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

Proportion:

$$Z = \frac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$$

- **TWO population**

Mean:

$$Z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$$

Proportion:

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$$

where $\hat{p}_p = \dfrac{x_1+x_2}{n_1+n_2}$

REGRESSION

- For the **least square regression equation,** $\hat{y} = b_0 + b_1 x$

  where $b_1 = \dfrac{SS_{xy}}{SS_{xx}}$ and $b_0 = \dfrac{1}{n}\left[\sum y - b_1 \sum x\right]$

- **Correlation coefficient,** r

  $$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

  where

  $$SS_{xx} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}, \qquad SS_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{n} \text{ and}$$

  $$SS_{xy} = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

- **Regression *t*-Test**

  $$t = \frac{b_1 - \beta_1}{s_b} \qquad \text{where } s_b = \frac{s_e}{\sqrt{SS_{xx}}} \qquad \text{and } s_e = \sqrt{\frac{SS_{yy} - b_1 SS_{xy}}{n-2}}$$